

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
6 September 2002 (06.09.2002)

PCT

(10) International Publication Number
WO 02/068579 A2

- (51) International Patent Classification⁷: **C12G**
- (21) International Application Number: PCT/US02/00284
- (22) International Filing Date: 10 January 2002 (10.01.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
09/756,696 10 January 2001 (10.01.2001) US
- (71) Applicant: **PE CORPORATION (NY)** [US/US]; 761
Main Avenue, Norwalk, CT 06859 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW.

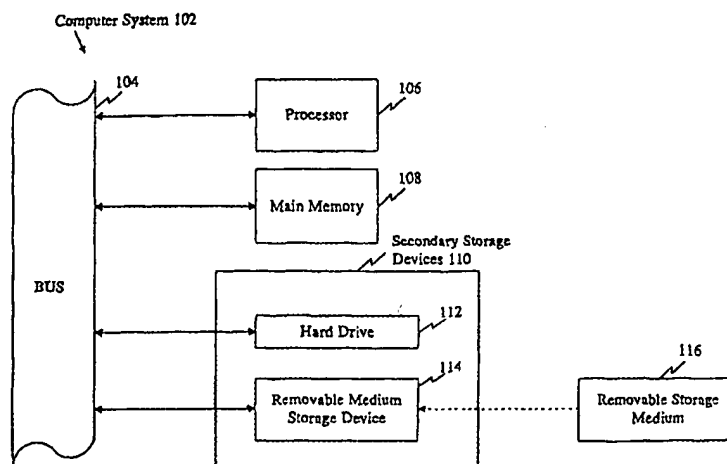
(84) Designated States (*regional*): ARIPO patent (GH, GM, KI, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

- (72) Inventors: **VENTER, Craig, J.**; Celera Genomics, 45 West Gude Drive C2-4#21, Rockville, MD 20850 (US). **ADAMS, Mark**; Celera Genomics, 45 West Gude Drive C2-4#21, Rockville, MD 20850 (US). **LI, Peter, W., D.**; Celera Genomics, 45 West Gude Drive C2-4#21, Rockville, MD 20850 (US). **MYERS, Eugene, W.**; Celera Genomics, 45 West Gude Drive C2-4#21, Rockville, MD 20850 (US).

Published:
— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

- (54) Title: **KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN EXONS OR TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF**



- (57) Abstract: The present invention is based on the sequencing and assembly of the human genome. The present invention provides the primary nucleotide sequence of the coding portion of the human genome in the form of a series of transcript sequences with accompanying exon information. This information can be used to generate nucleic acid detection reagents and kits such as nucleic acid arrays, and for other uses.

WO 02/068579 A2

CQ728483
 LOCUS CQ728483 884 bp DNA linear PAT 03-FEB-2004
 DEFINITION Sequence 14417 from Patent WO02068579.
 ACCESSION CQ728483
 VERSION CQ728483.1 GI:42297418
 KEYWORDS .
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 REFERENCE 1
 AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
 TITLE Kits, such as nucleic acid arrays, comprising a majority of
 humanexons or transcripts, for detecting expression and other uses
 thereof
 JOURNAL Patent: WO 02068579-A 14417 06-SEP-2002;
 PE Corporation (NY) (US)
 FEATURES Location/Qualifiers
 source 1. .884
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 9.4%; Score 298; DB 6; Length 884;
 Best Local Similarity 100.0%; Pred. No. 3e-159;
 Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 946 GAGGGCTATATCAAGTACTCTGCACTCTTCTATGGCTACTACAACAACCAGAGGACCATC 1005
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 187 GAGGGCTATATCAAGTACTCTGCACTCTTCTATGGCTACTACAACAACCAGAGGACCATC 246
 Qy 1006 GGGTGGCTGAGGTACCGGCTGCCTATGGCTTACTTTATGGTGGGGGTGAGCGTGTTTCGGC 1065
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 247 GGGTGGCTGAGGTACCGGCTGCCTATGGCTTACTTTATGGTGGGGGTGAGCGTGTTTCGGC 306
 Qy 1066 TACAGCCTGATTATTGTGTCATTCGATCGATGGCCAGCAATACCCAAGGAAGCACAGGCGAA 1125
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 307 TACAGCCTGATTATTGTGTCATTCGATCGATGGCCAGCAATACCCAAGGAAGCACAGGCGAA 366
 Qy 1126 GGGGAGAGTGACAACCTTCACATTCAGCTTCAAGATGTTTACCAGCTGGGACTACCTGATC 1185
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 367 GGGGAGAGTGACAACCTTCACATTCAGCTTCAAGATGTTTACCAGCTGGGACTACCTGATC 426
 Qy 1186 GGGAATTCAGAGACAGCTGATAACAAATATGCATCCATCACCACCAGCTTCAAGGAAT 1243
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 427 GGGAATTCAGAGACAGCTGATAACAAATATGCATCCATCACCACCAGCTTCAAGGAAT 484

CQ728483
 LOCUS CQ728483 884 bp DNA linear PAT 03-FEB-2004
 DEFINITION Sequence 14417 from Patent WO02068579.
 ACCESSION CQ728483
 VERSION CQ728483.1 GI:42297418
 KEYWORDS .
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 REFERENCE 1
 AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
 TITLE Kits, such as nucleic acid arrays, comprising a majority of
 humanexons or transcripts, for detecting expression and other uses
 thereof
 JOURNAL Patent: WO 02068579-A 14417 06-SEP-2002;
 PE Corporation (NY) (US)
 FEATURES Location/Qualifiers
 source 1. .884
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.:	4.1e-81	Length:	884
Score:	99.00	Matches:	99
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.93%	Indels:	0
DB:	6	Gaps:	0

US-10-792-307-4 (1-906) x CQ728483 (1-884)

Qy	311	GluGlyTyrIleLysTyrSerAlaLeuPheTyrGlyTyrTyrAsnAsnGlnArgThrIle	330
Db	187	GAGGGCTATATCAAGTACTCTGCACTCTTCTATGGCTACTACAACAACCAGAGGACCATC	246
Qy	331	GlyTrpLeuArgTyrArgLeuProMetAlaTyrPheMetValGlyValSerValPheGly	350
Db	247	GGGTGGCTGAGGTACCGCTGCCTATGGCTTACTTTATGGTGGGGGTCAGCGTGTTCGGC	306
Qy	351	TyrSerLeuIleIleValIleArgSerMetAlaSerAsnThrGlnGlySerThrGlyGlu	370
Db	307	TACAGCCTGATTATTGTTCATTCGATCGATGGCCAGCAATACCCAAGGAAGCACAGGCGAA	366
Qy	371	GlyGluSerAspAsnPheThrPheSerPheLysMetPheThrSerTrpAspTyrLeuIle	390
Db	367	GGGGAGAGTGACAACCTTCACATTCAGCTTCAAGATGTTCAACAGCTGGGACTACCTGATC	426
Qy	391	GlyAsnSerGluThrAlaAspAsnLysTyrAlaSerIleThrThrSerPheLysGlu	409
Db	427	GGGAATTCAGAGACAGCTGATAACAAATATGCATCCATCACCACCAGCTTCAAGGAA	483